Journal of Soft Computing in Civil Engineering 8-3 (2024) 128-154

Contents lists available at SCCE

Journal of Soft Computing in Civil Engineering

Journal homepage: www.jsoftcivil.com

Efficient Ensemble Learning-Based Models for Plastic Hinge Length Prediction of Reinforced Concrete Shear Walls

Naser Safaeian Hamzekolaei [1,*]; Mohammad Sadegh Barkhordari [2]

1. Assistant Professor, Department of Civil Engineering, Bozorgmehr University of Qaenat, Qaen, Iran
2. Ph.D., Department of Civil Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran
* Corresponding author: *nsafaeian@buqaen.ac.ir*

https://doi.org/10.22115/SCCE.2023.408189.1688

ARTICLE INFO

Article history:
Received: 22 July 2023
Revised: 27 October 2023
Accepted: 20 November 2023

Keywords:
Reinforced concrete wall (RCSW);
Plastic hinge length (PHL);
Machine learning;
Ensemble algorithms;
Gradient boosting regressor (GBR);
SHapley additive exPlanations (SHAPE).

ABSTRACT

Reinforced concrete shear wall (RCSW) significantly improves the seismic performance of buildings. Accurate estimation of the plastic hinge length (PHL) of RCSWs is crucial as it significantly impacts the plastic deformation, ultimate displacement, and ductility capacity of RCSWs. This study aims to develop practical machine-learning (ML) models for PHL prediction of RCSWs. For this purpose, 721 data of nonplanar and rectangular RCSWs were utilized. Deep neural network-based ensemble learning models including Simple Averaging Ensemble (SAE), Stacking Ensemble (SE), Snapshot Ensemble (SSE), and Deep Forest (DP), were leveraged. Meanwhile, inherently ensemble-learning-based (IELB) algorithms including the XGBoost, RandomForest, CatBoost, HistGradientBoosting, AdaBoost, Bagging, ExtraTrees, and GradientBoosting regressor, and data-driven empirical equations were considered for comparison. The Taylor diagram and statistical comparison of the results revealed that the proposed SE model with the gradient boosting regressor (GBR) meta-learner (MAE=0.043, MSE=0.012, R^2=0.916) outperformed all employed deep-and IELB-based ensemble algorithms as well as the empirical formulas for PHL of RCSWs. The SHapley Additive exPlanations-based model interpretation together with Sobol's sensitivity analysis results revealed that the wall length is the most crucial input variable, followed by the effective height and the axial load ratio.

How to cite this article: Safaeian Hamzehkolaei N, Barkhordari MS. Efficient ensemble learning-based models for plastic hinge length prediction of reinforced concrete shear walls. J Soft Comput Civ Eng 2024;8(3):128–154. https://doi.org/10.22115/scce.2023.408189.1688

2588-2872/ © 2024 The Authors. Published by Pouyan Press.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

In recent years, the application of machine learning (ML), as one of the most effective artificial intelligence (AI) approaches, for effectively modeling and forecasting structural performance has become popular [1,2]. Various ML algorithms have been established and employed to predict the structural design of buildings [3,4], the strength of material [5–7], collapse and damage identification [8,9], structural reliability analysis [10], interfacial-bond/axial/flexural/shear strength [1,11,12], and so on [13–15].

Chou et al. [16] developed metaheuristics-optimized single and ensemble machine learning models by using extreme gradient boosting (XGBoost), jellyfish search, and symbiotic organisms search algorithms for estimating the nominal shear capacity of the shear wall systems. Barkhordari and Massone [17] employed ensemble deep neural network models for predicting the failure mode of the shear walls. Barkhordari and Es-haghi [18] investigated the predictive capability of a hybrid approach combining artificial neural networks (ANN) and the Simulated Annealing (SA) method for forecasting the response of the reinforced concrete shear walls (RCSWs) subjected to strong ground vibrations. Zhang et al. [19] utilized various ML algorithms to predict failure modes, shear, and deformation capacity of the shear walls. Barkhordari and Massone [20] utilized ensemble techniques and hybrid intelligence algorithms to predict the shear strength of squat-reinforced concrete walls. Parsa and Naderpour [21] investigated the ML-based shear strength prediction of the RCSWs using a hybrid support vector regression algorithm improved by, Particle Swarm, Teaching-learning, and Harris Hawks Optimization algorithms. Table 1 summarizes the information on the application of the ML models on the RCSWs.

Table 1
Summary of the application of the ML models on the RCSWs.

Application	Description	Refs.
Numerical model optimization	ML-based traditional layered shell model optimization procedure was investigated for accurate finite element analysis of RCSW structures.	[4]
Predicting the shear strength	ML models was employed to predict the shear strength of RCSWs based on the geometry, material properties, and reinforcement details. This information can be used to design safe and more efficient structures.	[16,18,21–24]
Predicting ductility and energy dissipation	The ductility and energy dissipation capacity of RCSWs were predicted based on the geometrical and mechanical characteristics. The model can be used to design structures that are less likely to collapse suddenly under extreme events.	[19,25,26]
Failure mode identification	A crucial component of structural engineering, particularly in the context of seismic design and structural safety, is the identification of the failure mode for RCSWs. Engineers can evaluate the structural performance and make the appropriate design changes by accurately identifying the failure modes.	[17,27–29]
Seismic risk assessments	An accurate ML-based seismic risk assessments framework was proposed to predict a building's post-earthquake damage state using structural properties and ground motion intensity measures as model inputs.	[30]

Ensemble learning (EL) has gained increasing popularity in recent years [31,32]. Multiple systems and committee-based learning are alternative names for ensemble learning, which involves training and combining multiple learners to tackle a learning challenge. Therefore, it is used in this study to develop various predictive models for plastic hinge length (PHL) of reinforced concrete structural walls (RCSW). RCSWs are frequently used as lateral resisting systems to withstand the earthquake load in seismic regions. In order to assess and create cost-effective designs, it is crucial to understand the response and behavior of RCSWs, particularly given the growing popularity of performance-based design approaches for new structures. For this purpose, many numerical models have been presented to predict the response of the RCSWs [33–35]. In this type of analysis, the common assumption is that the plastic hinge length (l_{ph}) is related to wall length (l_w), for example $l_{ph}=0.5l_w$. PHL has an influence on determining ultimate strength and drift capacity. Therefore, accurate prediction of l_{ph} is important. Although $l_{ph}=0.5l_w$ was employed in the UBC-97 and more recently in NZS 3101:2006, however, previous studies indicate that approximating the PHL as $0.5l_w$ can provide an overestimation of l_{ph} [36–38]. Given that the straightforward definition of $l_{ph}=0.5l_w$ is being utilized for seismic assessment, this overestimation of l_{ph} may be problematic. Lu et al. [39] explored slender RCSWs with two layers of vertical reinforcement. The tested walls were utilized to investigate the impact of reinforcement ties, aspect ratio, and axial load effect. The experimental findings proved that the minimal vertical reinforcing restrictions in place are insufficient to cause the development of a significant number of secondary flexural cracks in the plastic hinge. They reported that l_{ph} values were discovered to be substantially lower than $0.5l_w$. Moreover, for non-rectangular (nonplanar) RCSWs, is not appropriate to assume that l_{ph} is proportionate to the length of the wall [36]. This issue is related to the significantly higher shear stress that occurs in non-rectangular RCSWs. It has been determined that the shear stress demand is a significant factor in determining l_{ph} of the RCSWs [38,40]. According to Massone and Alfaro [41], the distributions of curvature along the height of the wall are non-linear because of variations in steel-related characteristics, such as hardening and reinforcement ratio. They demonstrated how the final curvature is influenced by the degree of hardening and reinforcing. They stated that the outcomes are conservative with a significant dispersion of results when estimating the plastic hinge length using $l_{ph}=0.5l_w$, and the hinge model concentrates all curvature at the base of the wall, often leading to an underestimation of the curvature.

Various empirical equations have been proposed to estimate l_{ph} of RCSWs [36,38,42,43]. However, in general, empirical equations may have not an acceptable accuracy, since (1) the limited number of samples are used, (2) the limited range of input design variables are considered, (3) walls with non-rectangular sections may not considered for developing some equations, and (4) some of the formulas are established as a result of numerical or experimental research on beams or columns [36,42]. The literature review shows that there is a need for a new model considering a comprehensive database of the wall types and also effective parameters (e.g., shear stress) for the estimation of l_{ph} because PHL has a significant effect on the

calculation of plastic deformation, ultimate displacement, and ductility capacity of the RCSWs [41,42].

The main objective of this study is to evaluate the efficiency of the ensemble learning method-based ML frameworks for the PHL prediction of the RCSWs. For this purpose, ensemble learning algorithms including simple averaging ensemble (SAE), stacking ensemble (SE), and snap-shot ensemble (SSE) as well as deep forest (DP) algorithm are employed. The game theory-based interpretation approach is also utilized to explain the impact of the input variables on the outcomes of the best model. Meanwhile, inherently ensemble-learning-based (IELB) algorithms, including XGBoost, RandomForest, CatBoost, HistGradientBoosting, AdaBoost, Bagging, ExtraTrees, and GradientBoosting regressor, and existing mechanics-driven and data-driven equations are also investigated for comparison of the results.

The rest of the paper is organized as follows. The structural parameters of RCSW systems and the experimental dataset are presented in Section 2. The applied individual and ensemble models are discussed in Section 3. Performance evaluation of the applied ML models, the SHAP analysis and the sensitivity analysis to explain the impact of input variables are investigated in Section 4. Finally, concluding remarks are presented in Section 5.

2. Dataset for PHL of the RCSW

Fig. 1 shows the geometry and effective structural parameters of the common RCSW. Hoult [42] has created a brand-new dataset including 721 samples of the nonplanar and rectangular RCSWs, which is used in this work. The target variable is the PHL (l_{ph}) of the RCSWs. The secondary cracking ratio (r_{scr}), wall length (l_w), effective height ($H_e = M/V$), axial load ratio (ALR = P/P_n), and shear stress variable (SSV) are the inputs. The secondary cracking ratio and the shear stress variable can be calculated according to Eqs. (1) and (2), respectively.

$$r_{scr} = \frac{\rho_w}{\rho_{w,\min}} \le 1, \quad \rho_w = n_{sl}\frac{A_{sl}}{A}, \quad \rho_{w,\min} = \frac{(t_w - n_t d_{tr}) f_{cts}}{f_{ult} t_w} \tag{1}$$

$$SSV = \frac{V_{Max}}{A_g \sqrt{f_c'}} \tag{2}$$

where V_{Max}, A, f_c', f_{cts}, A_{sl}, n_{sl}, t_w, d_{tr}, f_{ult}, and ρ_w are the maximum shear force, the section area of the RCSWs, concrete compressive strength, concrete tensile strength, the section area of the longitudinal bar, number of bars in the longitudinal direction, wall thickness, transverse bars diameter, ultimate strength of the bars, and longitudinal reinforcement ratio, respectively.

Table 2 shows the characteristics of the database where Ave, Std, Min, and Max stand for average, standard deviation, minimum, and maximum, respectively. The histogram and correlation matrix of input variables are also presented in Fig. 2. It can be seen from Fig. 2 that

only the wall length (l_w) has a correlation coefficient higher than 0.5 with effective height ($H_e = M/V$). The shear stress variable (SSV) has a correlation coefficient of 0.45. All other input variables have a correlation coefficient less than 0.5 which indicates a weak relationship between them.

Table 2
The statistical data for input parameters.

Variable	Unite	Min.	Max.	Ave.	Std.
Secondary cracking ratio (r_{scr})	-	0.15	1.00	0.95	0.13
Wall length (l_w)	(m)	0.51	9.00	4.05	2.61
Effective height $(H_e = M/V)$	(m)	0.50	54.90	12.54	11.05
Axial load ratio (ALR)	-	0.00	0.43	0.09	0.07
Shear stress variable (SSV)	-	0.02	2.22	0.29	0.23
Plastic hinge length (PHL)	(m)	0.021	4.889	1.188	0.87

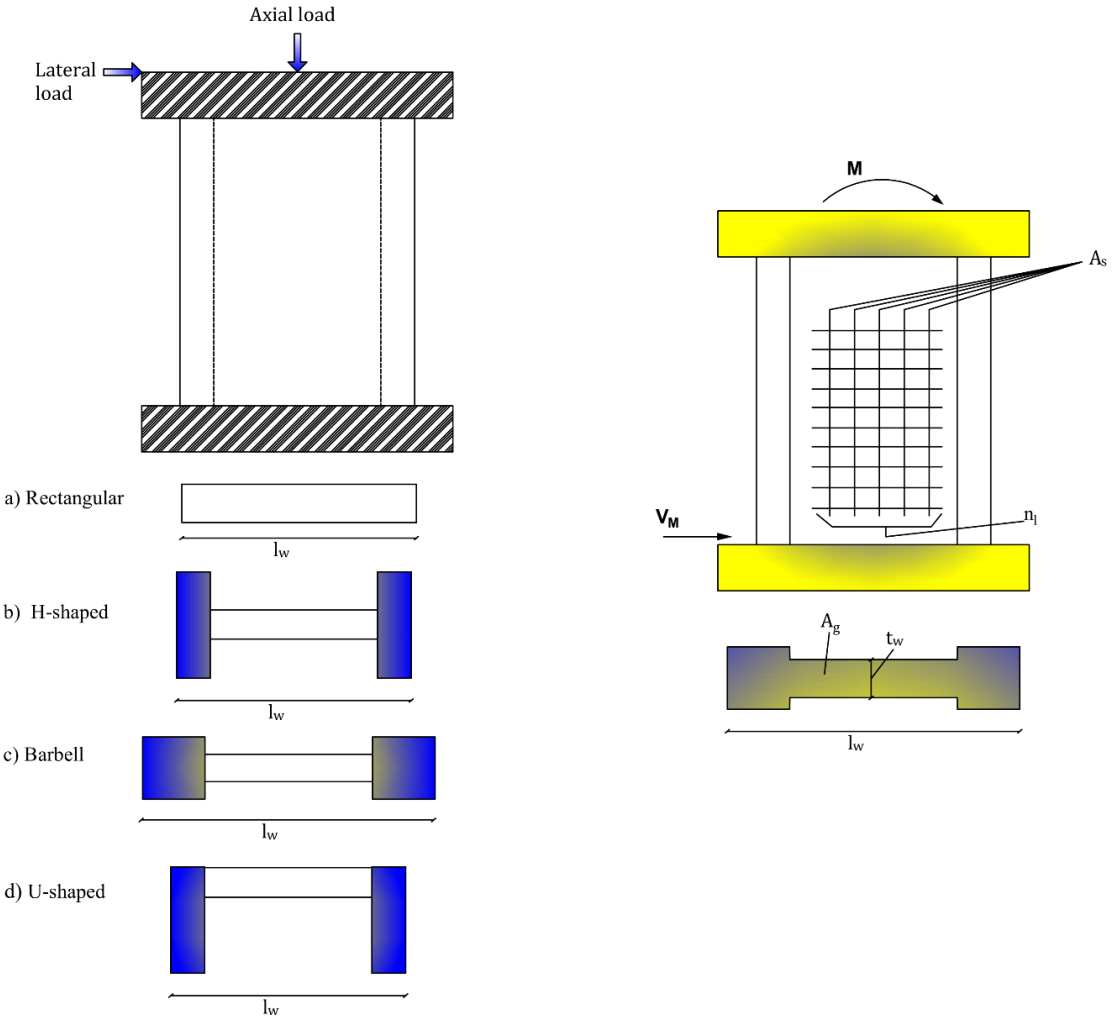

Fig. 1. Geometry, loadings, and sections of the RCSWs.

Following that, the data are randomly split into training and testing sets using the conventional split of 80% to 20%, respectively [5]. Then, the database is transformed by scaling each feature to a range [-1,1], as follows:

$$f_n = \frac{2(f - f_{\min})}{(f_{\max} - f_{\min})} - 1 \tag{3}$$

where f, f_n, $f_{\min}$, and $f_{\max}$ denote a feature of the sample, the normalized value, the minimum and maximum values of the feature, respectively.

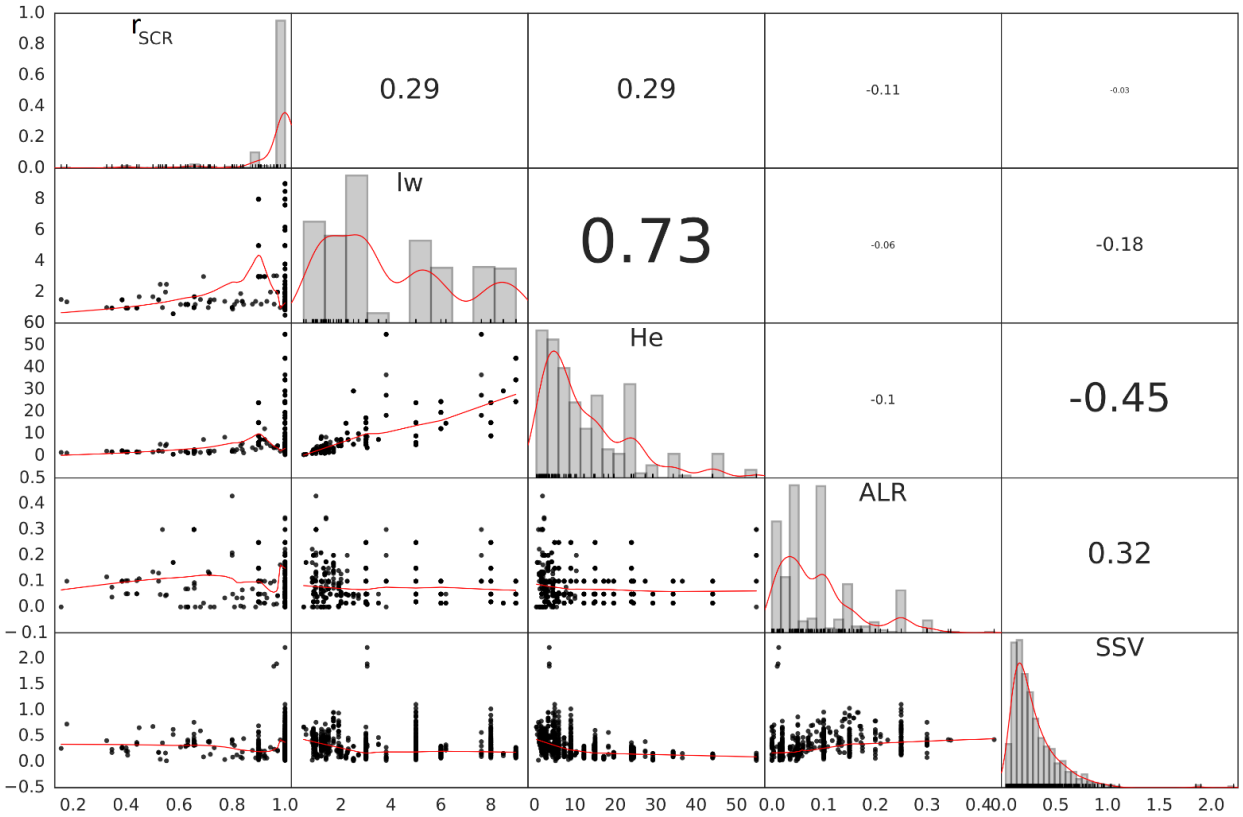

Fig. 2. Histogram and correlation matrix of inputs.

Fig. 3 shows the relationship between each input parameter and the l_{ph}. As can be seen, there is a negative correlation between the ALR and SSV versus l_{ph} as well as a positive correlation between the secondary cracking ratio (r_{scr}), wall length (l_w), and effective height ($H_e = M/V$) versus l_{ph}. Therefore, the calculation of the PHL based on the conventional equation ($l_{ph} = 0.5 l_w$) may not be enough accurate.

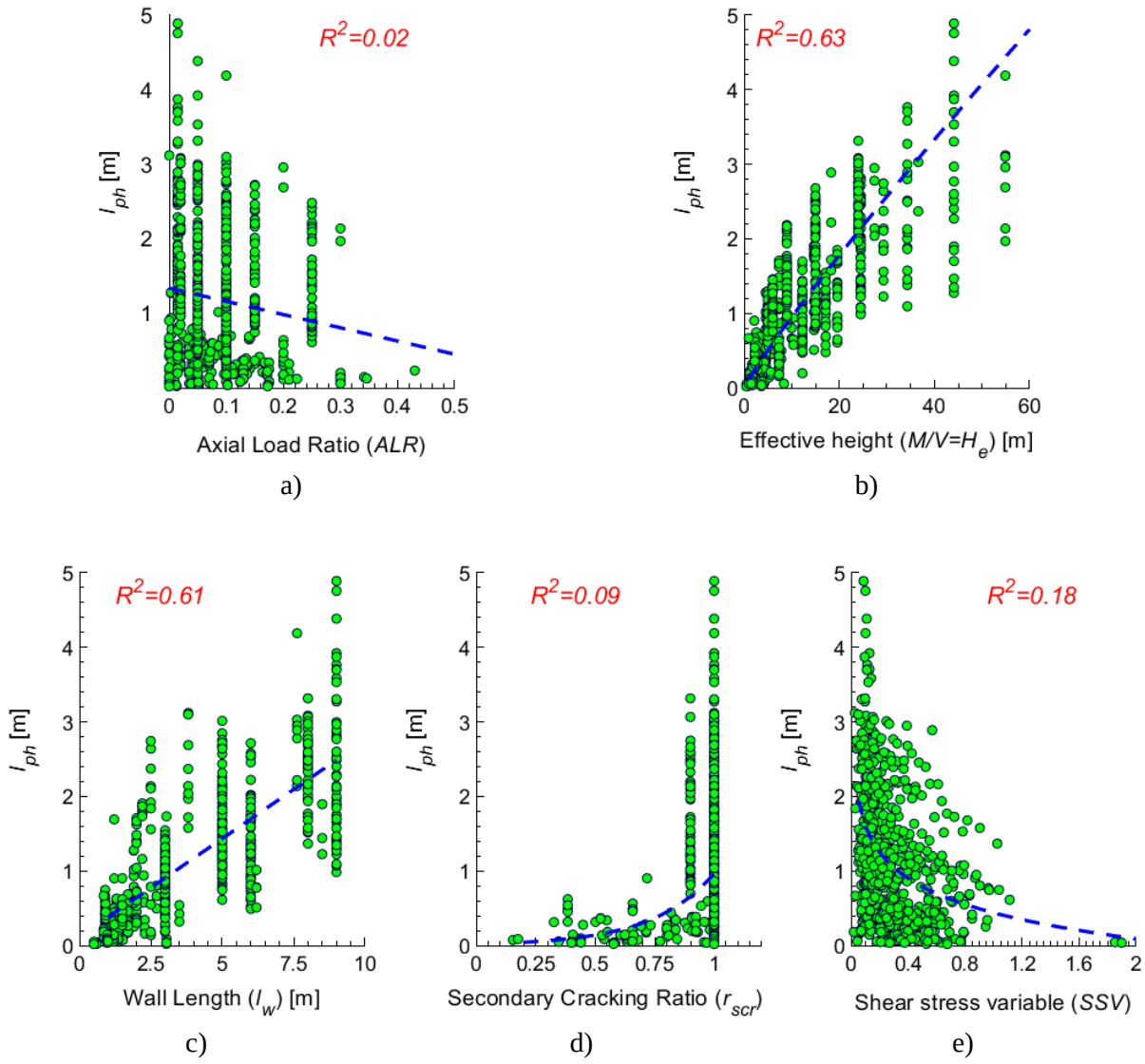

Fig. 3. Input and output parameters correlation.

3. Individual and ensemble models

In ensemble learning, firstly, a number of individual base models, called base learners, are trained separately utilizing common/arbitrary algorithms. Then, the base learners are merged with a unified strategy (Fig. 4). If all basic learners in an ensemble learning algorithm are of the same type, the algorithm is called homogeneous. In this study, homogeneous ensemble learning algorithms with either deep neural network (DNN) or decision trees (DT)-based base learners are employed. The applied models in this study include simple averaging ensemble (SAE), stacking ensemble (SE), snap-shot ensemble (SSE), and deep forest (DP) algorithm [44,45].

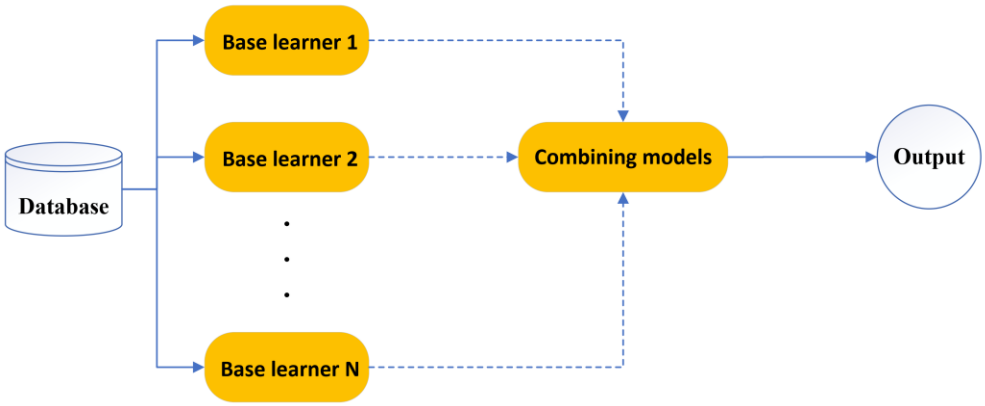

Fig. 4. General workflow of ensemble learning models.

3.1. Base learners

An artificial Neural Network (NN) [20] is composed of interconnected artificial neurons that resemble the neurons in a biological brain. When a neural network has more than one hidden layer with computational neurons, it is called a deep neural network (DNN) [20]. Co-adaptation is more likely to occur when there are many neurons in a fully connected layer. Co-adaptation happens when multiple neurons in a layer derive similar hidden characteristics from the inputs or highly similar hidden features. To address this issue, dropout can be used during NN training. Dropout algorithms randomly disable a portion of neurons in a layer during each training step. The dropout rate represents the proportion of neurons that are deactivated.

Table 3
Characteristics of the applied DDNs.

		Base-model number				
		1	**2**	**3**	**4**	**5**
Drop rate (DR) and layer ith neurons number (LiNN)	L1NN	35	15	15	10	15
	DR	0.01	0.01	-	0.01	0.01
	L2NN	25	15	15	20	15
	L3NN	-	45	20	15	30
	DR	-	-	-	0.01	0.01
	L4NN	-	-	35	20	25
	DR		-	0.01	-	-
	L5NN	-	-	-	15	25
	L6NN	-	-	-	-	35
Activation		Tanh	Tanh	Tanh	Tanh	Tanh
Optimizer		Adam	Adam	Adam	Adam	Adam
Coefficient of determination		0.87	0.85	0.87	0.89	0.90

In this study, the SAE, SE, and SSE utilize DNNs as base learners, while DP uses decision trees. Therefore, the first step is to select several DNNs with different architectures to be used in subsequent steps. For this purpose, five DNNs with different architectures are developed using training data. These trained DNNs are then utilized as base learners. The GridSearchCV algorithm [46] is employed to determine the activation function, optimizer, number of neurons, and dropout rate values for each DNN. The characteristics of the employed base learners are presented in Table 3.

3.2. Simple averaging ensemble (SAE) algorithm

During the prediction process, the deep learning model calculates each output value. An ensemble model is formed by combining the predictions of multiple DNN models to make a unified decision instead of relying on individual predictions [20]. This allows for simultaneous forecasting of each value in the data using multiple models. The performance of an ensemble model is measured by taking the average of the predicted values [20].

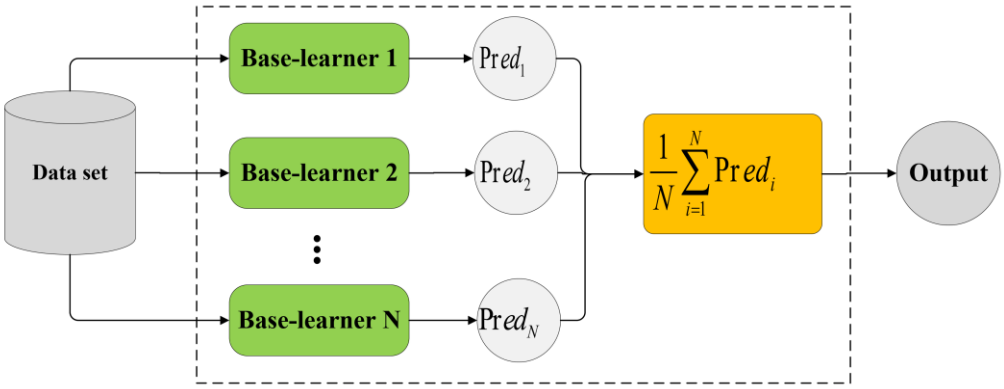

Fig. 5. Workflow of the SAE algorithm.

Fig. 5 shows the workflow of the SAE [20], where N represents the number of base learners and $Pred_i$ (i=1,..,N) represents the estimated value by base learners. It should be mentioned that the output layer has a single neuron and a linear activation function (Fig. 5).

3.3. Stacking ensemble (SE) algorithm

Stacking [20], also known as stacked generalization, refers to the process of studying a learning algorithm that combines the results of multiple learning algorithms. The main goal is to reduce generalization error by utilizing a pool of base learners and employing another learner to aggregate their estimates [20]. The stacking algorithm is an ensemble method with two operating levels: level 0 consists of the base learners, while level 1 includes a meta-model. Specifically, in this study, DNNs are used as the base learners at level 0, and their outputs serve as input data for the meta-models. The meta-models employed in this study are support vector regressor (SVR), decision tree regressor (DTR), gradient boosting regressor (GBR), random forest regressor (RFR), Ada boost regressor (ABR), and bagging regressor (BR). Fig. 6 illustrates the workflow of the SE algorithm, showing the interaction between the meta-model and the base learners.

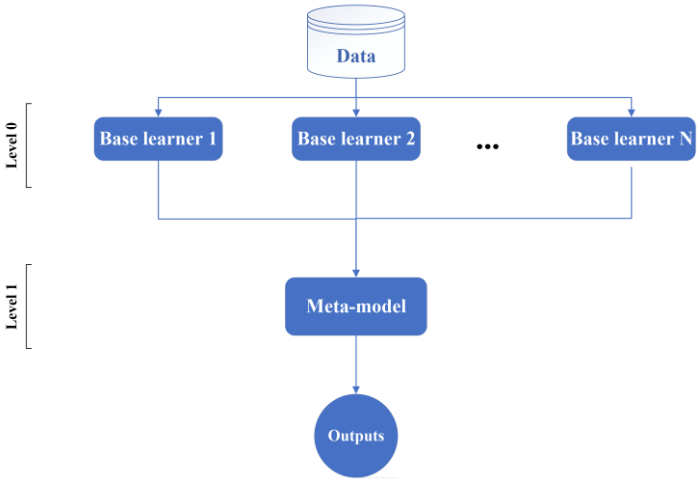

Fig. 6. Workflow of the SE algorithm.

3.4. Snap-shot ensemble (SSE) algorithm

The methods used for tuning DNNs are stochastic gradient descent (SGD) and its variations [20]. While avoiding erroneous solutions due to fictitious local minima is generally considered beneficial, some researchers argue that these local minima contain valuable data that can improve the model performance [20]. The SSE algorithm was designed to train a group of DNNs with an easy-to-implement training mechanism. It leverages the non-convex structure of DNNs and the capability of the SGD to escape from local minima. Instead of training M independent DNNs, the SSE algorithm allows SGD to converge M times to local minima along its optimization path. Each time the model converges, its weights are saved and added to a pool. The procedure then continues with a high learning rate to escape the current local minimum (Fig. 7).

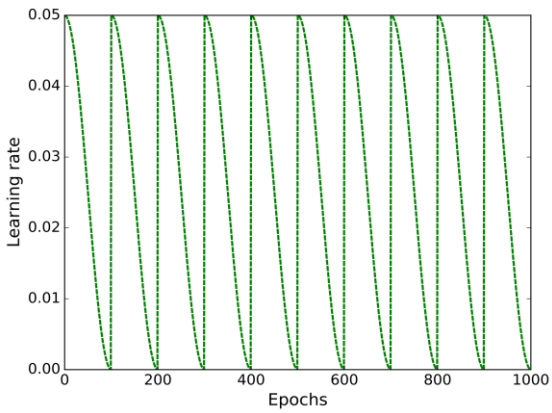

Fig. 7. Learning schedule of the SSE algorithm.

Specifically, the SSE technique employs a cycling process where the learning rate is abruptly increased and then rapidly decreased, following a cosine function. The resulting ensemble in the SSE method comprises snapshots of the optimization loop. Based on the accuracy, the structure of the base model 5 (Table 2) is selected to develop SSE. To study the impact of the number of members on accuracy, a sensitivity analysis is conducted. Since models saved towards the end of the session have undergone more training epochs compared to those saved earlier, it is possible

that they perform better. Therefore, the list of recorded models is reversed, and models from the pool are chosen and added starting from the last, thereby increasing the number of SSE members.

Fig. 8 illustrates the performance of SSE as the number of members increases from 2 to 20. Generally, the performance remains relatively stable until reaching 10 model members. However, as the number of members increases from 10 to 20, the performance of the proposed SSE model declines. This may be attributed to the fact that the first models generate outlier predictions due to their fewer training epochs and insufficient time to learn all the information. Furthermore, it is worth noting that even with just two members, the SSE model outperforms individual models. In this study, the SSE with 10 recorded models is utilized for PHL prediction of RCSW.

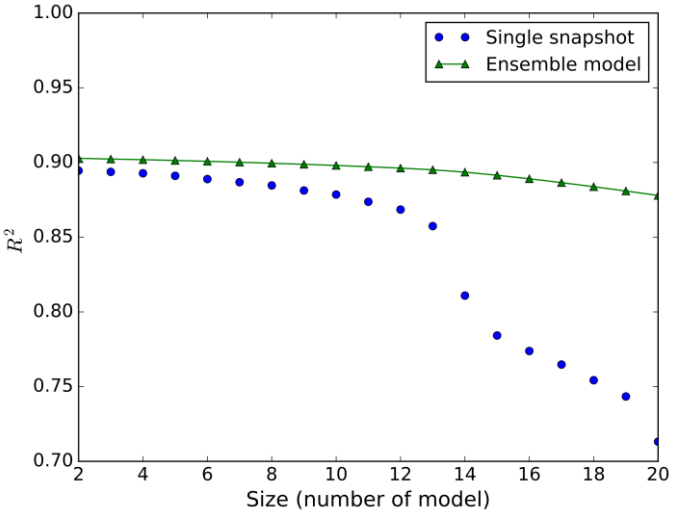

Fig. 8. The performance of the SSE model versus the size of the model.

3.5. Deep forest (DP) algorithm

Using the concept of Variation in Memory (VIM) of base learners, the DP algorithm aims to provide a model with a strong ability to make precise estimations. The DP algorithm incorporates various methods for VIM simultaneously, including sample manipulation, input feature manipulation, and output representation manipulation [44]. Deep Forest is a relatively new ensemble learning technique that utilizes decision trees and deep learning. Traditional decision tree algorithms construct a single tree to model the data, which may result in overfitting or under fitting. Deep forest can overcome this issue by constructing multiple levels of decision trees. The initial layer of trees is built using original data, and the subsequent layers are constructed using the results from the previous layer. This approach allows the creation of a hierarchical structure of decision trees that can uncover more complex correlations in the data. The ability of Deep Forest algorithm to handle noisy, high-dimensional data is one of its key advantages. High-dimensional data can cause overfitting or under fitting in standard decision tree techniques. To address this issue, Deep Forest employs a strategy where a random subset of features is chosen for each tree in every layer. This process helps reduce the impact of irrelevant or distracting features on model predictions. Additionally, each layer is utilized in a way similar to a random forest to further mitigate overfitting.

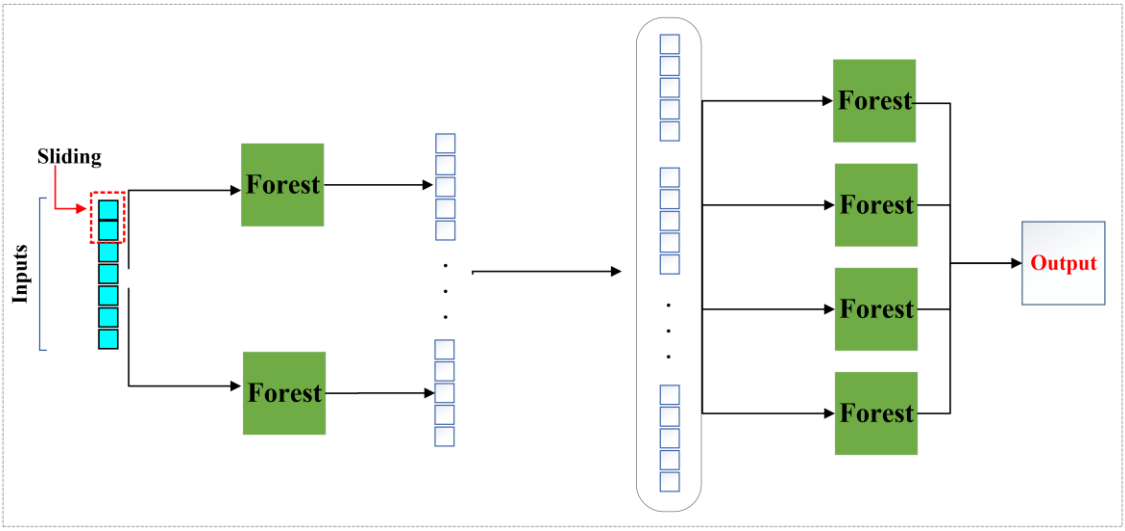

Fig. 9. Workflow of deep forest algorithm.

As shown in Fig. 9, a sliding window in DP algorithm slides over the features of the samples. Then, a set of decision trees are trained using the selected features. In this scenario, each tree is randomly trained with a restricted number of samples. The sliding windows generate a large number of outputs for each decision tree. The output of these decision trees serves as the input for the subsequent decision trees. This process is repeated until there is no further improvement in the accuracy of the model for the validation data. The key advantage of this algorithm is its adaptive structural complexity. Instead of predefining the number of layers, the algorithm dynamically increases the number of layers until the lowest error value is reached or no further improvement in accuracy is observed. In this study, the recommended default values have been used for the DP algorithm. These include 500 trees in each forest, a sliding window size of 1, and eight forests.

3.6. Performance metrics

In this study, the coefficient of determination (R^2), mean-squared error (MSE), and median absolute error (MAE) are utilized to evaluate the efficiency of the applied ML models. A higher R^2 value, along with lower MSE and MAE values, indicates a better fit.

$$MSE = \sum_{i=1}^{N} \left(\frac{(O_{vi} - E_{vi})^2}{N} \right) \tag{4}$$

$$R^2 = \frac{[\sum (O_{vi} - \bar{O}_v)(E_{vi} - \bar{E}_v)]^2}{\sum (O_{vi} - \bar{O}_v)^2 (E_{vi} - \bar{E}_v)^2} \tag{5}$$

$$MAE = median(|O_{v,1} - E_{v,1}|, ..., |O_{v,i} - E_{v,i}|) \tag{6}$$

where N is the number of samples, O_v denotes the observed value, $\bar{O}_v$ denotes the average of recorded values, E_v represents the estimated value, and $\bar{E}_v$ represents the average of estimated values.

4. Results and discussion

In this section, the efficiency and accuracy of the applied models (SAE, SSE, DF, SVR, DTR, GBR, RFR, ABR, and BR) are evaluated.

Table 4 presents the MAE, MSE, and R^2 values for the training and testing phases. The boxplots for the predicted PHL values are also presented in Fig. 10. Good agreement between the results of the applied models and the observed PHL values are quite obvious from Fig. 10. Based on the training results in Table 4, the SE-DTR is the best model with highest R^2 value and the lowest MAE and MSE values. SE-RFR ranks second among all models. Both SE-RFR and SE-BR have the same MSE and MAE values, with a slight difference in the R^2 values. Also, the SSE algorithm had the worst performance in the training stage in terms of MSE, MAE, and R^2 values. From the results of the testing phase, it is noticeable that there is inconsistency in the performance assessment results across the three indices. For instance, SE-RFR performs the best in terms of MAE, while SE-GBR exhibits the best performance in terms of MSE and R^2 among all investigated models. However, this is not the case for MAE. Moreover, the MAE value for the SE-GBR model is close to that of the SE-RFR model. Therefore, considering the overall performance, the SE-GBR model is selected as the best model for the subsequent steps of this study.

Table 4
Performance assessment of ensemble models.

Model		Training data			Testing data		
		MSE (m^2)	**R^2**	**MAE (m)**	**MSE (m^2)**	**R^2**	**MAE (m^2)**
SAE		0.011	0.911	0.043	0.015	0.898	0.046
SSE		0.012	0.898	0.047	0.021	0.858	0.068
DF		0.008	0.938	0.023	0.020	0.861	0.059
SE	SVR	0.009	0.926	0.045	0.014	0.900	0.053
	DTR	**0.0**	**1**	**0.0**	0.018	0.877	0.058
	GBR	0.003	0.978	0.028	**0.012**	**0.916**	0.043
	RFR	0.002	0.984	0.014	0.013	0.907	**0.042**
	ABR	0.008	0.934	0.056	0.017	0.883	0.079
	BR	0.002	0.980	0.014	0.015	0.898	0.049

The variations of the observed and predicted PHL values for the best model (SE-GBR) are presented in Fig. 11. It is obvious from Fig. 11 that the SE-based GBR model's predicted PHL values thoroughly follow the corresponding observed PHL values.

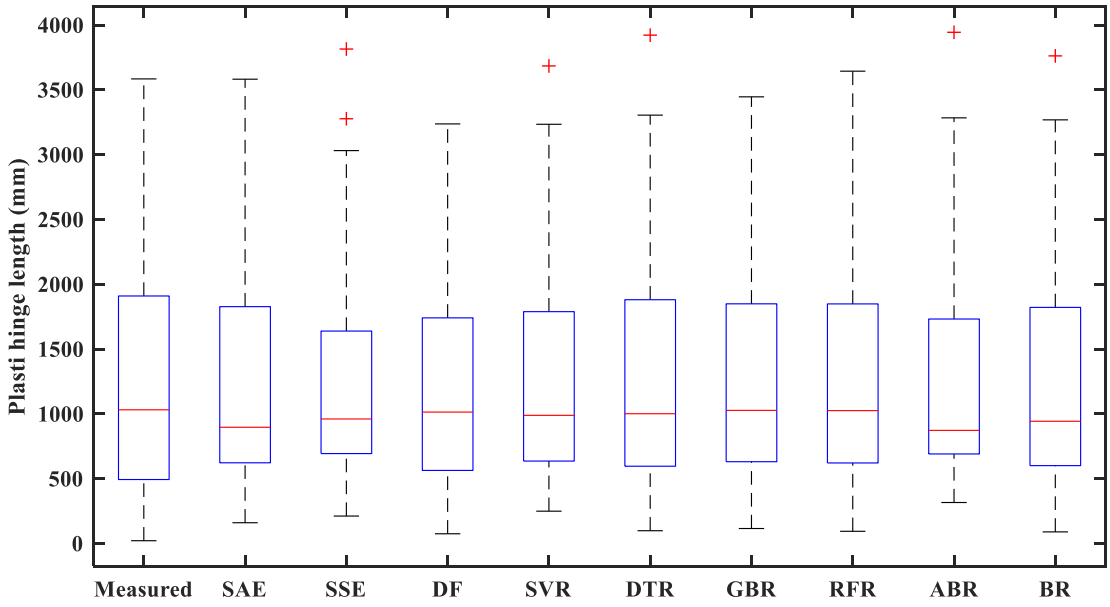

Fig. 10. Box plots of predicted PHL values (testing dataset).

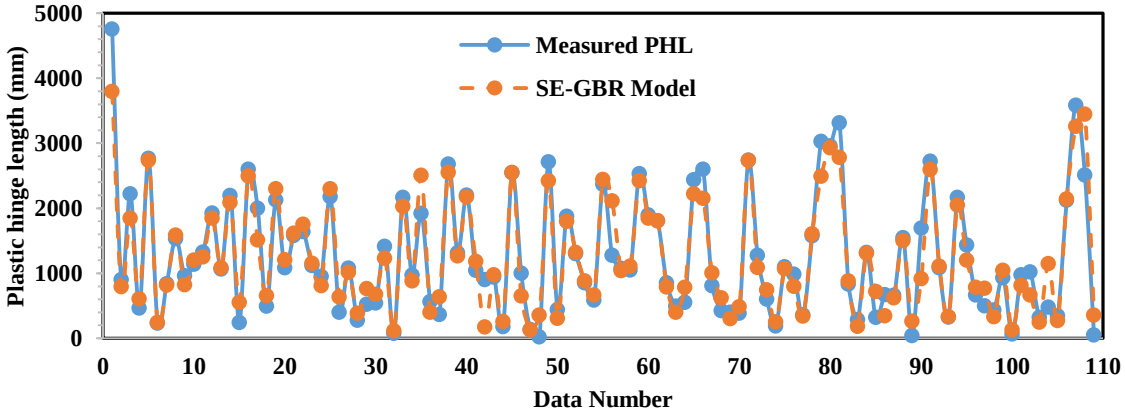

Fig. 11. Variation of the observed and predicted PHL values (testing dataset).

4.1. Inherently ensemble-learning algorithms

In this section, the performance of the proposed SE-GBR ensemble learning model is compared with inherently ensemble-learning-based (IELB) algorithms. For this purpose, eight different IELB algorithms including the XGBoost, RandomForest, CatBoost, HistGradientBoosting, AdaBoost, Bagging, ExtraTrees, and GradientBoosting regressor are investigated.

In Table 5, the results of the applied IELB algorithms and the proposed SE-GBR model are compared. As can be seen, the performance indices for the SE-GBR model are superior than those for the IELB algorithms. However, the MAE value for SE-GBR and ExtraTrees models is almost the same. It should be mentioned that the ExtraTrees algorithm (with MSE=0.016, R^2=0.89, and MAE=0.046) performs best among other IELB algorithms. However, the MSE and R^2 of the RandomForest are the same as ExtraTrees model. Therefore, the RandomForest algorithm (with MAE=0.5) has gained the second rank among all IELB algorithms. These

results, in general, verify that the proposed SE-GBR model outperformed the applied inherently ensemble-learning-based algorithms for PHL prediction of the RCSWs.

Table 5
Comparison of IELB algorithms and SE-GBR model (testing data).

Model	MSE (m^2)	R^2	MAE (m)
XGBoost	0.024	0.834	0.096
RandomForest	0.016	0.890	0.050
CatBoost	0.018	0.875	0.047
HistGradientBoosting	0.016	0.888	0.056
AdaBoost	0.028	0.809	0.101
Bagging	0.017	0.879	0.062
ExtraTrees	0.016	0.890	0.046
GradientBoost	0.017	0.880	0.081
SE-GBR	0.012	0.916	0.043

4.2. Comparison with empirical equations

This section provides a brief overview of the empirical equations derived from the literature for comparison purpose. For this purpose, six existing empirical models proposed in [36,38,42,43,47,48] are used. Bohl and Adebar [36] employed a nonlinear finite element (FE) model to evaluate the flexural deformation capacity of RCSWs. Through testing various RCSWs with different shear span ratios, lengths, and axial loads, they derived an equation that predicts the equivalent plastic hinge length, as presented in Eq. (10). Priestley et al. [43] proposed a plastic hinge formula to meet specific serviceability and damage limits. Kazaz [38], based on a regression study, developed a model that considers variables such as wall length, shear span ratio, axial load ratio, and reinforcement ratio. This model was calibrated using both the Turkish seismic code requirements and an FE shear wall model. Two different methods were employed to estimate the length of the plastic hinge zone. The first method involved evaluating the strain distribution across the wall's cross-section to determine the ultimate curvature configuration, whereas, the second method required calculating the curvature arrangement using strain data from shell elements applied to the FE model at the opposing corners of the wall. Furthermore, Hoult [42] introduced Eq. (12) for predicting the PHL of RCSWs. This equation was derived using standard regression techniques and based on the same database utilized in this study.

A list of the empirical equations for calculating the PHL of RCSWs is provided in Eqs. (7)-(12).

NZS 3101.1:2006, ASCE/SEI 41-06 [47]
$$l_{ph} = 0.5 l_w \tag{7}$$

Priestley et al. [43]
$$l_{ph} = \beta H_e + 0.2 l_w + 0.022 f^{ye} d_{bl}$$
$$H_e = M/V,\ \beta = 0.2(f^{ult}/f^{y} - 1) \le 0.08,\ f^{ye} = 1.1 f^{y} \quad (8)$$

Eurocode 8 [48]
$$l_{ph} = \frac{H_e}{30} + 0.2 l_w + \frac{0.11}{\sqrt{f_c^{'}}}(d_{bl} f^{y}) \quad (9)$$

Bohl and Adebar [36]
$$l_{ph} = (0.2 l_w + 0.05 H_e)(1 - 1.5\frac{P}{P_n}) \le 0.8 l_w \quad (10)$$

Kazaz [38]
$$l_{ph} = 0.27 l_w (1 - \frac{P}{P_n})(1 - \frac{f^{y}\rho_{tr}}{f_c^{'}})(\frac{H_e}{l_w})^{0.45} \quad (11)$$

Hoult [42]
$$l_{ph} = r_{scr}(0.13 l_w + 1.8 H_e^{0.6}) \times (1 - 0.3\frac{P}{P_n}) \times (1 - 0.13\log(SSV)) \quad (12)$$

where P/P_n denotes the axial load ratio, d_{bl} represents the bars diameter, f^{ult} denotes the ultimate strength of the bars, $f_c^{'}$ denotes the compressive strength of the concrete, f^{y} represents the yield strength of the bars, and ρ_{tr} represents the transverse reinforcement ratio.

The testing data was used for comparison between the accuracy of the empirical equations (7-12) and the proposed SE-GBR model. The predictions from the SE-GBR model are adjusted based on the range of the dataset using Eq. (3). The performance of the applied empirical equations is presented in Table 6 and Fig. 12. The boxplots of the estimated PHL values are also shown in Fig. 13. It is evident that the predictions of the PHL using empirical equations (7) and (9) deviated significantly from the dashed line for Fig. 12(a) and Fig. 12(c), respectively. Figs. 12(a), (b), and (e) demonstrate that Eqs. (7), (8), and (11) overestimates the PHL compared with observed values. It should be mentioned that previous studies have also reported that approximating the PHL by $0.5 l_w$ can overestimate the real values of the l_{ph} [42]. This overestimation can be attributed to the neglect of other influential factors. Moreover, from Fig. 12(c), it can be concluded that Eq. (9) underestimates the PHL (Fig. 13).

Table 6
Indicators value for SE-GBR and empirical models.

Model	Performance indices		
	MSE (m^2)	R^2	MAE (m)
Eq. (7)	1.292	0.30	0.668
Eq. (8)	0.297	0.65	0.303
Eq. (9)	0.405	0.53	0.258
Eq. (10)	0.235	0.73	0.250
Eq. (11)	0.492	0.43	0.194
Eq. (12)	0.200	0.77	0.206
SE-GBR	**0.012**	**0.916**	**0.043**

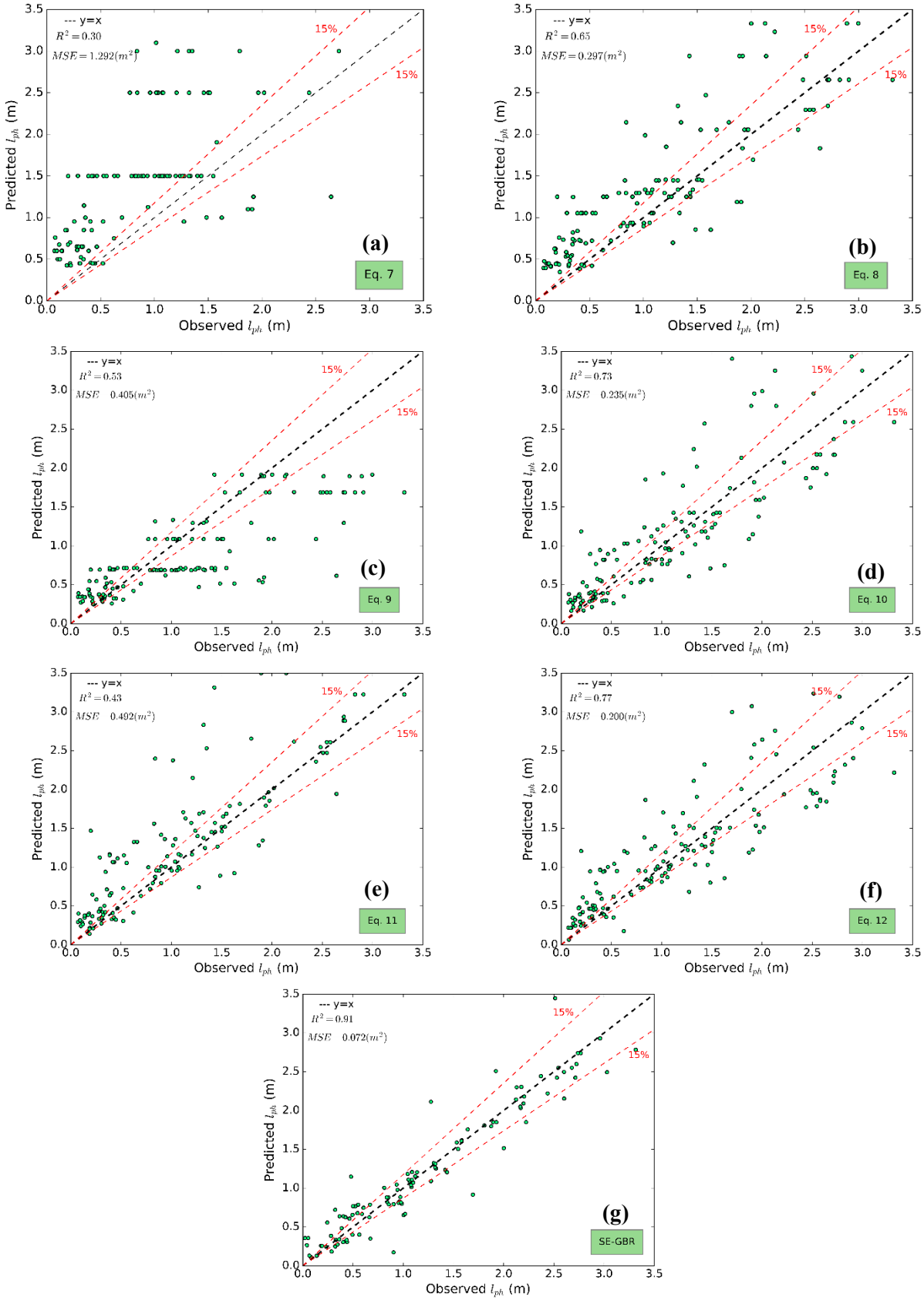

Fig. 12. Comparison of the results of existing equations Vs. SE-GBR model.

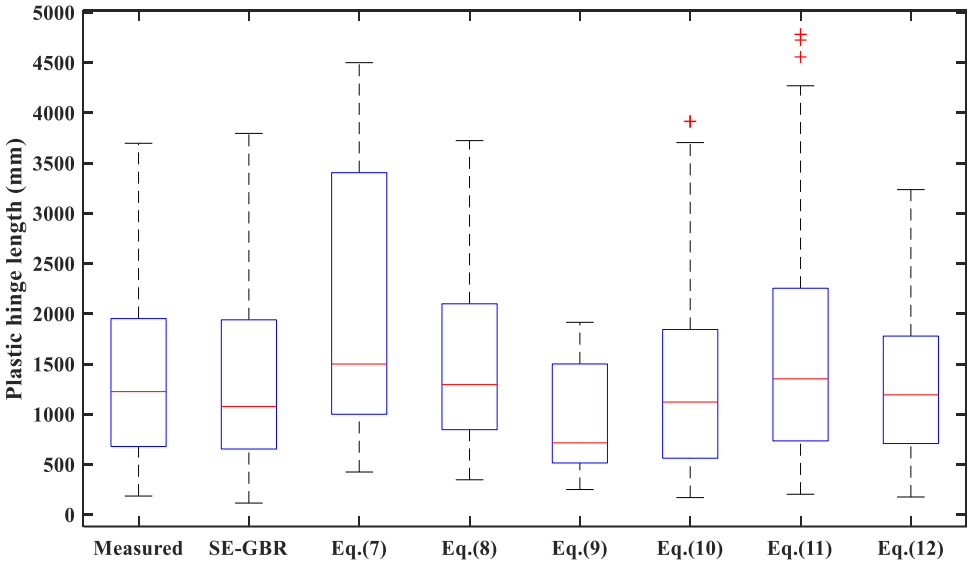

Fig. 13. Box plots of residual error (test data).

According to Table 6, Eq. (10) with R^2 of 0.73 and Eq. (12) with R^2 of 0.77 demonstrate better agreement with the experimental results compared to other empirical equations. Among the empirical equations, Eq. (12) with the MAE value of 0.206 m and the lowest MSE (0.2 m^2) has the best performance. Nevertheless, the proposed SE-GBR model, with higher R^2 values and lower MSE and MAE values, performed better than all existing formulas. The good agreement between the results of the SE-GBR and the observed PHL values is obvious from Fig. 13.

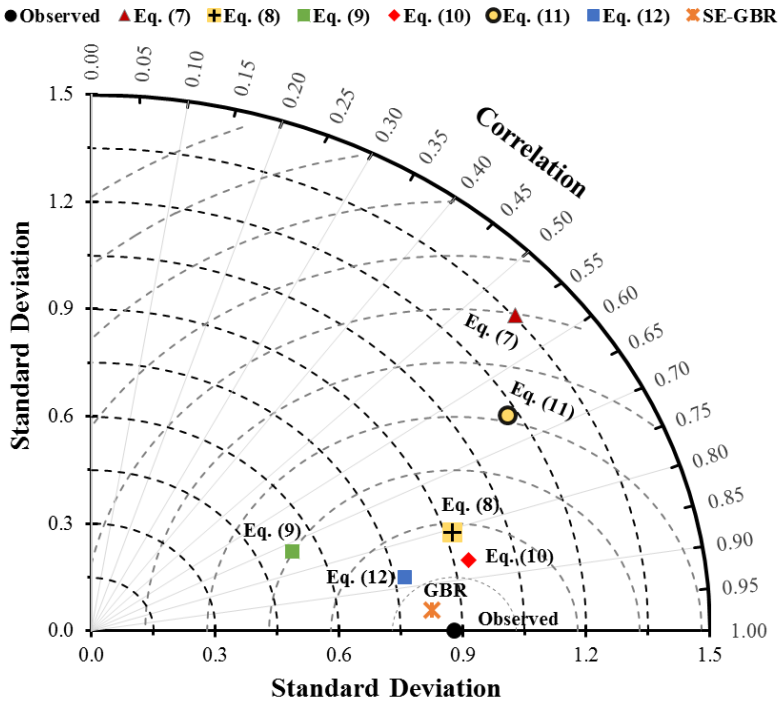

Fig. 14. Taylor diagram for the estimated PHL values.

The Taylor diagram for all applied models is depicted in Fig. 14. In this figure, the azimuth angle denotes to the correlation (r) between the predicted values and the observed PHL values, and the radial distance from the observed PHL values denotes the centered RMS error. As can be seen in Fig. 14, among all applied algorithms, the SE-GBR model, having the highest correlation with observations and the lowest RMS error (less than 0.1), agrees best with the observations values. Fig. 14 also indicates that Eq. (12) has a higher correlation with observations and a lower centered RMS error than other empirical equations. However, the SE-GBR has a lower distance from the observed PHL values compared with all empirical models. Therefore, it can be concluded that the proposed model provided the most accurate results.

4.3. Model explain-ability

It is crucial to understand the results of ML-based models and be able to interpret how and why these models provide predictions in a human-readable manner [49]. In this study, the state-of-the-art method known as SHAP (SHapley Additive exPlanations) [50] is employed to explain the impact of input variables on the prediction capability of the SE-GBR model. SHAP combines game theory concepts with linear surrogate modeling techniques and provides additional mathematical guarantees concerning the accuracy of the explanations. In this method, the average influence of a feature on the prediction through all possible subsets of features is denoted as the Shapley value [49].

The SHAP-based relevance analysis results of the input variables affecting PHL of RCSWs are presented in Fig. 15. This figure shows the global importance of each input feature on the output of the SE-GBR model. It should be mentioned that the feature importance of input variables is calculated as the mean value of the absolute Shapley values of the dataset. Fig. 15 demonstrates that the wall length (l_w) is the most crucial input variable affecting the PHL of the RCSWs, followed by the effective height (H_e=M/V) and the axial load ratio (ALR). Although SSV has an impact on how plasticity develops, it seems that this factor is not as important as other parameters.

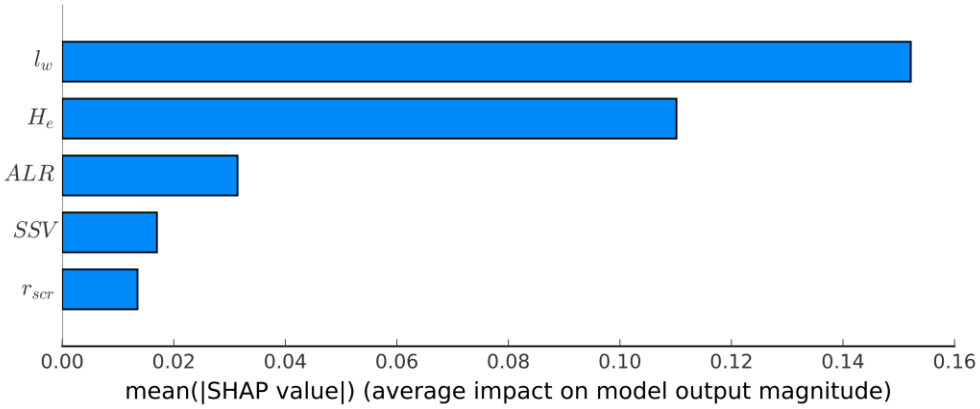

Fig. 15. Global importance of the input features.

Fig. 16 shows the summary plot for the SE-GBR model's prediction. In this figure, each point represents a Shapley value for a specific input variable and a single observation of the test set.

The color in this figure denotes the parameters' value, which varies from low (blue) to high (red). This figure shows the order of significance of input parameters affecting the PHL of RCSWs, i.e., whether the prediction is changed positively or negatively. As can be seen in Fig. 16, an increase in the wall length (l_w) increases the plastic hinge length of the RCSWs. This is due to the connection between tension shift effects and l_w. The tension shift action increases and the plasticity spreads farther with a long wall length. Moreover, the model is more likely to estimate larger l_w as the value of the effective height (M/V) increases, which is related to the lower nonlinear shear deformations and a higher possibility of flexural yielding before reaching the nominal shear strength of the wall. Moreover, it can be seen that the PHL has an inverse relationship with the axial load level (ALR). In the other words, increasing the axial load reduces PHL values which may be due to the simultaneous concrete crushing and reinforcement buckling in the compression area, or concrete crushing in the compression area prior to reinforcement yielding in the tension area. This caused a dramatic decrease of l_{ph} in axially loaded RCSWs by closing the distance between the bending moment at the yielding point and the ultimate capacity.

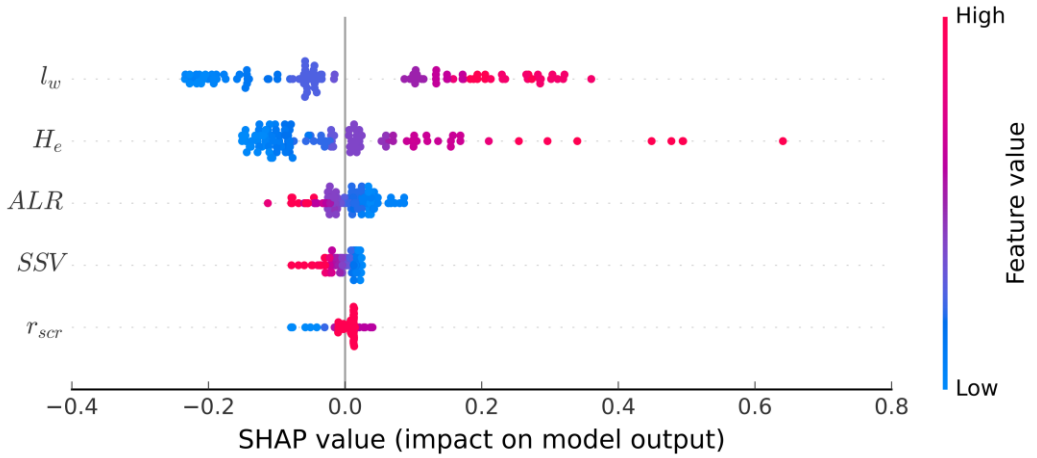

Fig. 16. Summary plot for SHAP analysis using SE-GBR model.

In addition, this analysis reveals that the higher values of the shear stress variable (SSV) will decrease l_{ph}. This issue can be related to (1) the interaction between nonlinear flexural and shear modes in RCSWs, and (2) the effects of the nonlinear shear deformations on the axial compressive strains even in slender RCSWs, which lead to reduced strength and ductility. It is also revealed that increasing the secondary cracking ratio (r_{scr}) increases the l_{ph}. This is in line with other studies that discovered a threshold amount of longitudinal steel is necessary to permit secondary cracking in RCSW system. The PHL will be shortened if the longitudinal reinforcement ratio of the wall is lower than the minimum value needed to permit secondary cracking.

4.4. Sensitivity analysis using Sobol's technique

In this section, the Sensitivity Analysis (SA) is performed to assess how a model's outcomes respond to variations in input variables. Sobol's SA method has gained considerable attention due

to its accuracy and simplicity [51,52]. This method is based on the decomposition of the variance of the model output into summands of variances of the input variables in increasing dimensionality. In this study, the variance-based Sobol's technique is employed for SA. Sobol's method takes a probabilistic perspective, treating the model's inputs as a multidimensional vector with a joint probability density function. The stochastic representation of the model can be expressed as follows:

$$Y = MO(x) \tag{13}$$

where MO represents a mathematical operator, x is a vector of random inputs, and Y denotes the quantity of interest. The Sobol's index for a subset 'u' of the input vector and total Sobol indices are defined as follows [51,52]:

$$S_i = Var(MO_i(x_i)) / Var(Y),\ S_i^{total} = \sum_{i=1}^{n} S_i \tag{14}$$

where $S_i\,(i = 1,..,n)$ denotes the Sobol's index, and S_i^{total} represents the total Sobol index.

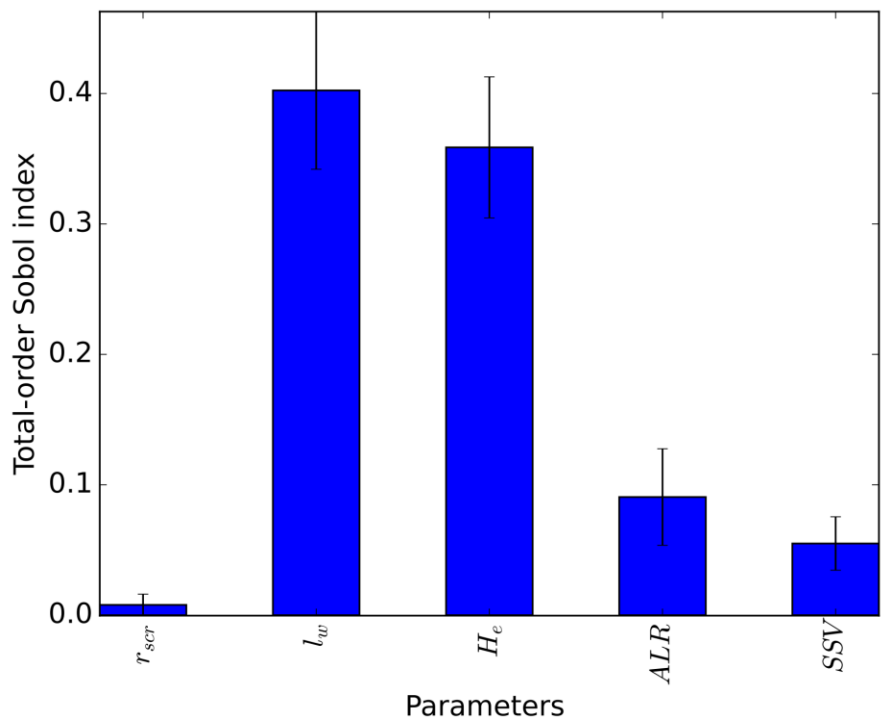

Fig. 17. Summary plot for SHAP analysis using SE-GBR model.

Fig. 17 shows the Sobol's technique-based sensitivity analysis results for PHL of the RCSW system. The SA results illustrate that the wall length and the wall effective height exhibit a higher total-order sensitivity coefficient. Consequently, they exert a significant influence on the PHL of the RCSWs. It is worth mentioning that, similar to the SHAP-based model explanations analysis, discussed in section 4.3, the SA results indicates that the secondary cracking ratio (r_{scr}), axial load ratio (ALR) have small effects and the shear stress variable (SSV) has negligible effects on the PHL of RCSWs.

5. Conclusion

In this study, ensemble machine learning (ML) algorithms including the simple averaging ensemble (SAE), stacking ensemble (SE), snap-shot ensemble (SSE), and deep forest (DP) algorithm were employed for prediction of the plastic hinge length (PHL) of reinforced concrete shear walls (RCSW) systems. For this purpose, a comprehensive dataset with 721 nonplanar and rectangular RCSWs was used. The secondary cracking ratio (r_{scr}), wall length (l_w), effective height ($H_e = M/V$), axial load ratio (ALR=P/P_n), and shear stress variable (SSV) are considered as input variables. The result of the proposed ML-based models was also compared with inherently ensemble-learning-based (IELB) algorithms including the XGBoost, RandomForest, CatBoost, HistGradientBoosting, AdaBoost, Bagging, ExtraTrees, and GradientBoosting regressor, as well as with available empirical equations for PHL of RCSWs. The performance of the applied models was investigated in terms of the regression plot, boxplot, Taylor diagram, and three statistical indices, i.e., median absolute error (MAE), mean-squared error (MSE), and coefficient of determination (R^2). The SHAP analysis, an acronym for SHapley Additive exPlanations, was also leveraged to interpret the degree of influence of each input variable on the prediction. The following conclusions can be expressed in light of the results of comparisons analysis:

- The stacking ensemble (SE)-based gradient boosting regressor (GBR) meta-learner with the statistical values of MAE=0.043 m, MSE=0.012 m^2, and R^2=0.916 outperformed other ensemble-based ML algorithms in the testing phase. Also, SE with a random forest regressor meta-model achieved the second rank among all other applied models.
- SE-GBR model outperformed all employed IELB algorithms. Moreover, between IELB algorithms, the ExtraTrees regressor with the statistical values of MAE=0.046 m, MSE=0.016 m^2, and R^2=0.89 performed better than all applied IELB algorithms.
- The proposed stacking ensemble-based gradient boosting regressor model surpassed all available empirical formulas for prediction of the PHL of RCSWs in terms of MSE, R^2, and MAE.
- Among the six available empirical equations for PHL prediction of RCSWs, the Bohl and Adebar's as well as Hoult's equations had better agreement with experimental results than other equations.
- The results of sensitivity analysis and the SHAP interpretation-based feature importance analysis revealed that the wall length (l_w) is the most influential input variable affecting the PHL of the RCSWs, followed by the effective height ($H_e = M/V$) and the axial load ratio (P/P_n).
- SHAP analysis results showed that an increase in the wall length, the effective height, and the secondary cracking ratio increases the plastic hinge length of RCSWs. In contrast, an increase in the axial load ration and shear stress decreases the plastic hinge length of the RCSWs.

Results of this study verified that the developed SE-based GBR model outperform all empirical equations and other compared IELB algorithms for estimating the PHL of RCSWs. However, due to data limitations, like other machine learning algorithms, the robustness and accuracy of the model output may be questionable for new big data. Expanding the data sources with new high-quality data, employing advanced feature-extracting techniques to identify additional relevant features, and combining machine learning models with advanced optimization algorithms could enhance the robustness and generalization capability of the predictive model for PHL of RCSWs by leveraging the strengths of different approaches.

Nomenclature

l_{ph}	Plastic hinge length	l_w	Wall length
ALR	Axial load ratio	$H_e = M / V$	effective height
f^{ult}	bars ultimate strength	$f_c^{'}$	strength of the concrete
f^{y}	bars yield strength	ρ_{tr}	transverse reinforcement ratio
A_{sl}	section area of longitudinal bar	SSV	shear stress variable
A	section area of the wall	t_w	wall thickness
n_{sl}	bars number in longitudinal direction	f_{ult}	ultimate strength of the bars
f_{cts}	concrete tensile strength	f_c'	concrete compressive strength
V_{Max}	maximum shear	SE	stacking ensemble
SAE	simple averaging ensemble	SSE	snap-shot ensemble
DP	deep forest		

Data availability

The entire database is available on the online repository Zenodo (https://zenodo.org/record/4321917).

Acknowledgments

The authors would like to acknowledge the financial support of Bozorgmehr University of Qaenat for this research under contract number 39246.

Funding

This research received no external funding.

Conflicts of interest

The authors declare that they have no conflict of interest.

Authors contribution statement

NSH and MSB: Data creation, Methodology, Software, Validation, Writing–original draft. NSH: Supervision, analysis of the results, Writing – review & editing.

References

[1] Raza A, El Ouni MH, Baili J, uz Zaman Khan Q. Data-driven analysis on axial strength of GFRP-NSC columns based on practical artificial neural network tool. Compos Struct 2022;291:115598. https://doi.org/10.1016/J.COMPSTRUCT.2022.115598.

[2] Isleem HF, Tayeh BA, Alaloul WS, Musarat MA, Raza A. Artificial Neural Network (ANN) and Finite Element (FEM) Models for GFRP-Reinforced Concrete Columns under Axial Compression. Mater 2021, Vol 14, Page 7172 2021;14:7172. https://doi.org/10.3390/MA14237172.

[3] Pizarro PN, Massone LM. Structural design of reinforced concrete buildings based on deep neural networks. Eng Struct 2021;241:112377. https://doi.org/10.1016/j.engstruct.2021.112377.

[4] Tao MX, Wang YL, Zhao JZ, Wang C. Optimisation of layered shell model for analysis of reinforced concrete shear walls based on machine learning. J Build Eng 2023;77:107434. https://doi.org/10.1016/J.JOBE.2023.107434.

[5] Barkhordari M, Armaghani D, Mohammed A, Ulrikh D. Data-Driven Compressive Strength Prediction of Fly Ash Concrete Using Ensemble Learner Algorithms. Buildings 2022;12:132. https://doi.org/10.3390/buildings12020132.

[6] Ashrafian A, Safaeian Hamzehkolaei N, Dwijendra NKA, Yazdani M. An Evolutionary Neuro-Fuzzy-Based Approach to Estimate the Compressive Strength of Eco-Friendly Concrete Containing Recycled Construction Wastes. Buildings 2022;12:1280. https://doi.org/10.3390/buildings12081280.

[7] Zhang G, Safaeian Hamzehkolaei N, Rashnoozadeh H, Band SS, Mosavi A. Reliability assessment of compressive and splitting tensile strength prediction of roller compacted concrete pavement: introducing MARS-GOA-MCS. Int J Pavement Eng 2022;23:5030–47. https://doi.org/10.1080/10298436.2021.1990920.

[8] Sadegh Barkhordari M, Jahed Armaghani D, G. Asteris P. Structural Damage Identification Using Ensemble Deep Convolutional Neural Network Models. Comput Model Eng Sci 2023;134:835–55. https://doi.org/10.32604/cmes.2022.020840.

[9] Hwang SH, Mangalathu S, Shin J, Jeon JS. Machine learning-based approaches for seismic demand and collapse of ductile reinforced concrete building frames. J Build Eng 2021;34:101905. https://doi.org/10.1016/j.jobe.2020.101905.

[10] Vatani A, Jafari-Asl J, Ohadi S, Safaeian Hamzehkolaei N, Ahmadabadi SA, Correia JAFO. An efficient random forest-based subset simulation method for reliability analysis of the marine structure piles subject to scour. Proc Inst Civ Eng Marit Eng 2023:1–42. https://doi.org/10.1680/jmaen.2022.020.

[11] Barkhordari MS, Armaghani DJ, Sabri MMS, Ulrikh DV, Ahmad M. The efficiency of hybrid intelligent models in predicting fiber-reinforced polymer concrete interfacial-bond strength. Materials (Basel) 2022;15:3019.

[12] Barkhordari MS, Feng DC, Tehranizadeh M. Efficiency of Hybrid Algorithms for Estimating the Shear Strength of Deep Reinforced Concrete Beams. Period Polytech Civ Eng 2022;66:398–410. https://doi.org/10.3311/PPci.19323.

[13] Siahkali MZ, Ghaderi A, Bahrpeyma A, Rashki M, Safaeian Hamzehkolaei N. Estimating Pier Scour Depth : Comparison of Empirical Formulations with ANNs, GMDH, MARS, and Kriging. J AI Data Min 2021;9:109–28. https://doi.org/10.22044/jadm.2020.10085.2147.

[14] Safaeian Hamzehkolaei N, Alizamir M. Performance Evaluation of Machine Learning Algorithms for Seismic Retrofit Cost Estimation Using Structural Parameters. J Soft Comput Civ Eng 2021;5:32–57. https://doi.org/10.22115/scce.2021.284630.1312.

[15] Abdellatif S, Raza A. Machine learning model for predicting ultimate capacity of FRP-reinforced normal strength concrete structural elements. Struct Eng Mech 2023;85:315–35. https://doi.org/10.12989/sem.2023.85.3.315.

[16] Chou J-S, Liu C-Y, Prayogo H, Khasani RR, Gho D, Lalitan GG. Predicting nominal shear capacity of reinforced concrete wall in building by metaheuristics-optimized machine learning. J Build Eng 2022;61:105046. https://doi.org/10.1016/j.jobe.2022.105046.

[17] Barkhordari MS, Massone LM. Failure Mode Detection of Reinforced Concrete Shear Walls Using Ensemble Deep Neural Networks. Int J Concr Struct Mater 2022;16. https://doi.org/10.1186/s40069-022-00522-y.

[18] Barkhordari MS, Es-Haghi MS. Straightforward prediction for responses of the concrete shear wall buildings subject to ground motions using machine learning algorithms. Int J Eng Trans A Basics 2021;34:1586–601. https://doi.org/10.5829/IJE.2021.34.07A.04.

[19] Zhang H, Cheng X, Li Y, Du X. Prediction of failure modes, strength, and deformation capacity of RC shear walls through machine learning. J Build Eng 2022;50:104145. https://doi.org/10.1016/j.jobe.2022.104145.

[20] Barkhordari MS, Massone LM. Ensemble techniques and hybrid intelligence algorithms for shear strength prediction of squat reinforced concrete walls. 1 2023;8:37. https://doi.org/10.12989/ACD.2023.8.1.039.

[21] Parsa P, Naderpour H. Shear strength estimation of reinforced concrete walls using support vector regression improved by Teaching–learning-based optimization, Particle Swarm optimization, and Harris Hawks Optimization algorithms. J Build Eng 2021;44:102593. https://doi.org/10.1016/J.JOBE.2021.102593.

[22] Naderpour H, Sharei M, Fakharian P, Heravi MA. Shear Strength Prediction of Reinforced Concrete Shear Wall Using ANN, GMDH-NN and GEP. J Soft Comput Civ Eng 2022;6:66–87. https://doi.org/10.22115/SCCE.2022.283486.1308.

[23] Barkhordari MS, Tehranizadeh M. Response estimation of reinforced concrete shear walls using artificial neural network and simulated annealing algorithm. Structures 2021;34:1155–68. https://doi.org/10.1016/j.istruc.2021.08.053.

[24] Nguyen DD, Tran VL, Ha DH, Nguyen VQ, Lee TH. A machine learning-based formulation for predicting shear capacity of squat flanged RC walls. Structures 2021;29:1734–47. https://doi.org/10.1016/J.ISTRUC.2020.12.054.

[25] Topaloglu B, Taskin Kaya G, Sutcu F, Deger ZT. Machine learning-based estimation of energy dissipation capacity of RC shear walls. Structures 2022;45:343–52. https://doi.org/10.1016/J.ISTRUC.2022.08.114.

[26] Yaghoubi ST, Deger ZT, Taskin G, Sutcu F. Machine learning-based predictive models for equivalent damping ratio of RC shear walls. Bull Earthq Eng 2023;21:293–318. https://doi.org/10.1007/s10518-022-01533-6.

[27] Mangalathu S, Jang H, Hwang SH, Jeon JS. Data-driven machine-learning-based seismic failure mode identification of reinforced concrete shear walls. Eng Struct 2020;208:110331. https://doi.org/10.1016/J.ENGSTRUCT.2020.110331.

[28] Xiong Q, Xiong H, Kong Q, Ni X, Li Y, Yuan C. Machine learning-driven seismic failure mode identification of reinforced concrete shear walls based on PCA feature extraction. Structures 2022;44:1429–42. https://doi.org/10.1016/J.ISTRUC.2022.08.089.

[29] Rao P, Pollayi H, Rao M. Machine learning based design of reinforced concrete shear walls subjected to earthquakes. J Phys Conf Ser 2022;2327:012068. https://doi.org/10.1088/1742-6596/2327/1/012068.

[30] Kourehpaz P, Hutt CM. Machine Learning for Enhanced Regional Seismic Risk Assessments. J Struct Eng 2022;148:04022126. https://doi.org/10.1061/(ASCE)ST.1943-541X.0003421.

[31] Feng DC, Liu ZT, Wang XD, Jiang ZM, Liang SX. Failure mode classification and bearing capacity prediction for reinforced concrete columns based on ensemble machine learning algorithm. Adv Eng Informatics 2020;45:101126. https://doi.org/10.1016/j.aei.2020.101126.

[32] Cakiroglu C, Bekdaş G. Predictive Modeling of Recycled Aggregate Concrete Beam Shear Strength Using Explainable Ensemble Learning Methods. Sustainability 2023;15:4957. https://doi.org/10.3390/su15064957.

[33] López CN, Massone LM, Kolozvari K. Validation of an efficient shear-flexure interaction model for planar reinforced concrete walls. Eng Struct 2022;252:113590. https://doi.org/10.1016/j.engstruct.2021.113590.

[34] Massone LM, López CN, Kolozvari K. Formulation of an efficient shear-flexure interaction model for planar reinforced concrete walls. Eng Struct 2021;243:112680. https://doi.org/10.1016/j.engstruct.2021.112680.

[35] Barkhordari MS, Tehranizadeh M, Scott MH. Numerical modelling strategy for predicting the response of reinforced concrete walls using Timoshenko theory. Mag Concr Res 2021;73:988–1010. https://doi.org/10.1680/jmacr.19.00542.

[36] Bohl A, Adebar P. Plastic hinge lengths in high-rise concrete shear walls. ACI Struct Journal-American Concr Inst 2012;109:123.

[37] Hoult R, Goldsworthy H, Lumantarna E. Plastic Hinge Length for Lightly Reinforced Rectangular Concrete Walls. J Earthq Eng 2018;22:1447–78. https://doi.org/10.1080/13632469.2017.1286619.

[38] Kazaz İ. Analytical Study on Plastic Hinge Length of Structural Walls. J Struct Eng 2013;139:1938–50. https://doi.org/10.1061/(asce)st.1943-541x.0000770.

[39] Lu Y, Henry RS, Gultom R, Ma QT. Cyclic Testing of Reinforced Concrete Walls with Distributed Minimum Vertical Reinforcement. J Struct Eng 2017;143:04016225. https://doi.org/10.1061/(asce)st.1943-541x.0001723.

[40] Hoult RD, Goldsworthy HM, Lumantarna E. Plastic Hinge Length for Lightly Reinforced C-Shaped Concrete Walls. J Earthq Eng 2020;24:1083–114. https://doi.org/10.1080/13632469.2018.1453419.

[41] Massone LM, Alfaro JI. Displacement and curvature estimation for the design of reinforced concrete slender walls. Struct Des Tall Spec Build 2016;25:823–41. https://doi.org/10.1002/tal.1285.

[42] Hoult R. Universal plastic hinge length for reinforced concrete walls. ACI Struct J 2022;119. https://doi.org/https://doi.org/10.14359/51734650.

[43] Priestley MJN, Calvi GM (Gian M, Kowalsky, Mervyn J.Powell GH. Displacement-Based Seismic Design of Structures. Earthq Spectra 2008;24:555–7. https://doi.org/10.1193/1.2932170.

[44] Zhou Z-H, Feng J. Deep Forest. Natl Sci Rev 2017;6:74–86. https://doi.org/10.1093/nsr/nwy108.

[45] Kunapuli G. Ensemble Methods for Machine Learning. New York: Simon and Schuster; 2023.

[46] Pedregosa F., Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, et al. Scikit-learn: Machine Learning in Python. J Mach Learn Res 2011;12:2825–30. https://doi.org/10.5555/1953048.2078195.

[47] NZS3101, "Concrete structures standard." Standards New Zealand. Wellington, NZ: 2006.

[48] Code P. Eurocode 8: Design of structures for earthquake resistance—part 3: Assessment and retrofitting of buildings. Inc Corrigendum March 2005.

[49] Barkhordari MS, Billah AHMM. Efficiency of Data-Driven Hybrid Algorithms for Steel-Column Base Connection Failure Mode Detection. Pract Period Struct Des Constr 2023;28:04022061. https://doi.org/10.1061/(ASCE)SC.1943-5576.0000741.

[50] Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Adv. Neural Inf. Process. Syst., vol. 2017- Decem, Neural information processing systems foundation; 2017, p. 4766–75.

[51] Ionescu-Bujor M, Cacuci DG. A Comparative Review of Sensitivity and Uncertainty Analysis of Large-Scale Systems—I: Deterministic Methods. Nucl Sci Eng 2004;147:189–203. https://doi.org/10.13182/NSE03-105CR.

[52] Cacuci DG, Ionescu-Bujor M. A Comparative Review of Sensitivity and Uncertainty Analysis of Large-Scale Systems—II: Statistical Methods. Nucl Sci Eng 2004;147:204–17. https://doi.org/10.13182/04-54CR.